DOCUMENT RESUME

ED 144 977 TM 006 508

AUTHOR Fremer, John
TITLE Setting and Evaluating Competency Standards for Awarding High School Diplomas.
NOTE 24p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New York, New York, April 5-7, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Academic Standards; *Evaluation Criteria; Evaluation Methods; *Graduation Requirements; *Guidelines; High School Students; *Policy Formation; School Policy; *Senior High Schools; Student Evaluation

ABSTRACT

Seven guidelines for setting minimum competency standards for high school graduation are presented: (1) Competency standards for the diploma should be developed by a process that gives major attention to the needs of students yet recognizes an institutions's responsibility to societal needs. (2) Competency standards should be developed using good management procedures. (3) Standards must be set at realistic levels. (4) Standards should be administered with an adequate student advisory and guidance system providing preparation and feedback. (5) Standards should be applied consistently within programs, across programs, and over time. (6) Standards should be backed fully by the home institution to increase the likelihood that the resulting diploma will be acceptable to other institutions. (7) Standards should be administered so that the students involved feel they have been well served. Three factors which are believed to influence a school district's actions are the external requirements of other institutions and agencies, the basic educational objectives of the school district, and the personal development goals of individual students. Seven possible assessment procedures are outlined. They are: development of competence scales, use of the traditional methods, review of test questions, identification of prerequisite skills, use of normative data, comparison of test scores to course performance, and use of preselected groups. (Author/MV)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED144977

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John Fremer, Educational Testing Service

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Setting and Evaluating Competency Standards for Awarding High School Diplomas[1]

John Fremer

Educational Testing Service

BACKGROUND

The idea of setting competency standards for high school graduates has captured the attention of a wide variety of groups throughout the nation. The most noticeable of these, of course, is the press. The February 28, 1977 issue of Time magazine, under the heading "Devalued Diploma" indicates that: ". . . in the past year, 'minimal competency testing' has become the hottest new catch phrase in public education."

As someone who has tried to keep up with developments in the area of minimal competency testing, my files are bulging with "news" of this sort. The following headlines taken from one newspaper, The New York Times, provide some indication of the way the press is treating this development:

- "Basic-Skill Tests Pushed for High School Diploma," March 14, 1976
- "Minimum-Competency Tests for Students, An Old Idea, Now Revived and Spreading," January 31, 1977
- "Requiring Test to Get a Diploma Brings Wide Attention to Denver," January 31, 1977
- "Certificates of High School Proficiency Stir Controversy," March 2, 1977

---

[1]Paper presented at the annual meeting of the National Council on Measurement in Education, New York City, April 1977. [Part of Symposium--Credits, Theories and Standards: New Applications for the Secondary Diploma]

TM006 508

Educational journals and related publications have also given the topic a great deal of attention. Some titles of articles and commentaries help illustrate some of the issues that have been considered:

"The Politics of Back-to-Basics," Joan Baum, Change, November 1976 (pp. 32-36).

"Changing High School Requirements," National Education Association, Briefing, June 1976, No. 11.

Competency Tests and Graduation Requirements, James P. Clark and Scott D. Thomson, National Association of Secondary School Principals, 1975.

So much has been published on the topic of minimum-competency testing as a diploma requirement that I will not attempt to review all that is available. Three sources, however, warrant mentioning. The first is the National Association of Secondary School Principals (NASSP). This organization has given the most comprehensive attention that I have seen to the issue of competency requirements and has gone on record as favoring competency testing as one component of high school graduation requirements.

The second source is the Legislative Review, a publication of the Education Commission of the States. The editor, Chris Pipho, has done a conscientious job of collecting information about state legislation and regulations regarding minimal competency testing. Mr. Pipho drew on this experience to give the following words of advice to educators and legislators at the National Conference on Minimum Competencies in New York on March 4, 1977:

> To educators--Don't underestimate the concern of the man in the street or of state legislators regarding the skills of high school graduates. The competency requirement movement is not the work of a few conservative legislators.

To legislators--Don't underestimate the technical problems associated with legislation that calls for competency requirements testing.

The third source I will mention is a document that I am using as a handout for this meeting. It is an Educational Testing Service Information Report entitled Basic Skills Assessment Around the Nation. It provides an overview of many activities at the state and district level related to basic skills testing as a diploma requirement or, in some cases, a grade-promotion requirement.

ISSUES RAISED BY THE COMPETENCY REQUIREMENTS MOVEMENT

Other observers of the minimum competency movement have raised a number of issues. Among them are these:

Meaning of the Diploma

There seems to be widespread agreement that a high school diploma should be a guarantee of at least some minimal level of skill development on the part of the high school graduate. Reference is often made to the fact that you can only get a "GED" (High School Equivalency Diploma) by passing the Tests of General Educational Development. Similar comments are being made about the California High School Proficiency Examination.

Implications for Curriculum

Representatives of professional associations, including the NCTE, NCTM, IRA, and NEA, have voiced concern that minimal comptency testing will tend to restrict the content of the high school program. A good statement of this position is contained in the December 1976 document "Minimal Competencies and Measures of Competence" prepared by Alan Purves

of the University of Illinois.

Implication for Students

What about the student who fails to meet competency standards for the high school diploma? There is clear agreement that assessments should occur early in the secondary school years and that remedial programs should be available to students who fail to meet standards. But will this result in tracking of students into such programs and out of the mainstream? What about students who never meet standards? What happens to them? Does the school's responsibility end when students have completed four years beyond grade eight? When the students reach a specified age?

Implications for Measurement Professionals

Should schools develop their own tests or use existing ones? What areas should be assessed and with what types of tests? Can machine-scoreable tests alone do the job? How should test-based and other information be combined to make decisions about individual students?

The question of setting and evaluating standards, therefore, has to be viewed as only one part of a much larger set of issues related to the introduction of competency standards. It is an important issue, though, with complexities that seemingly have not been recognized by many who have urged the adoption of competency requirements for the diploma.

GENERAL PROPERTIES OF GOOD SYSTEMS FOR SETTING STANDARDS

This section of my paper looks at the overall system of setting, applying, and studying the effects of standards. Here is an overview of seven guidelines or principles that are reviewed:

1. Ethics/Responsibility. Competency standards for the diploma should be developed by a process that gives major attention to the needs of students but which recognizes an institutions's responsibility to societal needs.

2. Management Standards. Competency standards should be developed using good management standards or procedures.

3. Realism. Standards must be set at realistic levels.

4. Adequacy of Preparation and Feedback. Standards should be administered with an adequate student advisory and guidance system so that high standards are not accompanied by high failure rates.

5. Consistency. Standards should be applied consistently within programs, across programs, and over time.

6. Acceptability/Transferability. Standards should be backed fully by the home institution to increase the likelihood that the resulting diploma will be acceptable to other institutions.

7. Student Satisfaction. Standards should be administered so that the students involved feel they have been well served.

Ethics/Responsibility

The issue of ethical responsibility is frequently ignored when educational systems and policies are evaluated. One simply assumes that this issue is part of any such evaluation. However, this factor should be explicitly identified as a significant one. A commitment to meet the educational needs of students should characterize all aspects of the standard-setting process. Adherence to the guidelines summarized in this paper would constitute one type of evidence that a school or program was meeting its responsibilities. Each institution, though, must analyze its own situation to determine the extent to which such constraints as limited time and money are dealt with in ways that adequately serve students.

A commitment to meeting the educational needs of students does not, of course, eliminate the need for schools to consider the implications for

society of decisions made about individual students. As students satisfy degree requirements they move toward the receipt of educational credentials. Keeton (1974, p. 3) makes the following observation about such credentials: "Credentials are power. Those who set the rules of the credentialing game are establishing the routes of access to that power."

As institutions set the rules of the game, they must balance the needs of students and those of society. Whenever the completion of a particular program will carry the expectation that a student has acquired certain competencies, an institution's evaluation procedures and standards must be adequate to certify the existence of such competencies.

Management Standards

Turning first to the process by which criterion standards are initially established, it will be useful to ask whether an institution has followed reasonable management standards. Management standards apply to the steps taken in the design and implementation of competency standards for the diploma, so that necessary decisions are made by qualified people in a manner that provides for consistent quality. In individual high school courses where performance standards are set primarily by the teacher, management standards apply to the selection of appropriate faculty members. Additional management standards relate to the evaluation of that faculty member's performance by his or her principal, department chairman, and perhaps by the students. Since the responsibility for establishing and applying competency standards for the diploma will not typically be placed on a single individual, it will be necessary to develop procedures that take into account the nature and variety of contributing individuals.

An important first step in evaluating the adequacy of an institution's

approach to standard-setting procedure is to determine whether overall responsibility for this task has been defined. The particular individual(s) who hold this responsibility can be expected to vary depending on the nature of the program, department, division, or other part of the institution within which evaluations take place. Although someone with training and experience in the field of evaluation would be best for the job, those without such background are often selected. In such instances, the use of trained evaluation consultants will be essential.

Even when overall responsibility for establishing performance standards can be placed in the hands of a skilled evaluator, many other individuals should contribute to the standard-setting process. For this purpose, it will be useful to seek assistance not only from faculty and students but also from members of the larger community.

The job of designing and developing competency standards cannot be handled adequately if decisions about particular students must be made under great time pressure. Providing adequate calendar time, and also adequate staff for the development of competency standards, will make it possible to plan and carry out pilot projects to try out standards setting systems. Procedures can be established for monitoring the application of standards and for following up on the students afterwards. Most institutions will find it useful to adopt an iterative process for developing competency standards. The best available sources of information should be used when setting standards for the first group of students tested. The experience of conducting the initial evaluations, though, should become part of the background for future evaluations.

John Valley of Educational Testing Service has suggested the use of an

outside "audit" as part of an evaluation of the adequacy of an institution's management standards for setting competency standards. The notion is borrowed from the area of financial management. A well-managed business concern checks periodically the adequacy of its accounting-bookkeeping procedures by calling for an outside audit. The audit process consists of tracing financial transactions from a variety of starting points using different routes, different systems, and different people. The goal is that of determining whether the same results can be obtained as were produced by the company's basic system. The application of the audit model to the assessment of student learning against standards for the diploma would require the periodic use of an independent assessment method, followed by a comparison of assessment outcomes.

The foregoing guidelines for good standard-setting procedures relate primarily to the planning and managerial skills of the participating staff. There are other criteria that can be applied to the system they design.

Realism

Critics of U.S. high schools often criticize the lack of rigor associated with the evaluation of students' learning experiences. Little attention has been paid to the reverse problem--setting unrealistically high standards. This problem is most likely to occur when the individuals participating in the setting or application of standards do not have sufficient direct knowledge of the performance of typical students. It is the responsibility of the institution to judge what is reasonable to expect of students before making decisions regarding competency standards.

High school faculty and consultants from other institutions or from the larger community should work together in the planning and developing

of assessment procedures. All parties concerned will need to collect and review evidence of what the average student can reasonably expect to learn from each of the programs offered. Interviews with students and reviews of student papers or products will be helpful ways of getting an idea of what is realistic to expect of students.

Adequacy of Preparation and Feedback

The student who is sufficiently prepared for an evaluation against diploma standards should be able to anticipate the result. The extent to which results will be predictable will depend on the quality and availability of information about standards and on the effectiveness of the student-advising process. The assessment system and the standards employed can be considered unfair to students if they lead to high rates of failure for students who have worked conscientiously over a period of time. Since the focus of competency standards should be on learning outcomes, not on time spent or activities engaged in, this criterion may seem inappropriate or paradoxical. Why should it be considered unfair if a student has failed to meet a particular competency standard after working hard? What is unfair is that the institution has not detected that the student was not learning enough and has not acted accordingly.

As institutions gradually develop their assessment and standard-setting policies, they must pay attention to the consequences of a student's failure to meet a specified performance standard. These consequences can range from minor inconveniences and disappointment to major setbacks and disruptions. For example, if a student fails to meet a particular diploma standard when tested in the ninth grade, his or her failure may be of little moment. The implications of failure, however, can be quite different when the student is

already in the last year of high school.

The reverse side of the issue of fairness, of course, is that of quality control. If award of the diploma is not to be an automatic consequence of completing course and attendance requirements, what steps can be taken both to maintain quality standards and still make failure a rare occurrence? Some form of ongoing student evaluation and feedback will best serve this need. Testing needs to take place early enough so that remedial programs can be planned and carried out. Students can then sit for follow-up examinations after they have further developed their skills and competencies.

Consistency

In implementing established competency standards, school districts will have to achieve a high degree of consistency within and among school programs and across time. What can an institution do to promote consistency in the application of standards? One strategy is to conduct periodic reviews of the results of applying a particular set of assessment procedures and standards. If institutional guidelines for applying standards are not being followed, a review of background materials and guidelines by participating staff will be needed.

It should be recognized that absolute consistency of assessment can never be attained. Each school district will need to consider the level of inconsistency that is "tolerable." Clearly the consequences of incorrect judgments will need to be considered and a balance struck among such possible competing factors as student needs and societal needs.

Acceptability/Transferability

Although arrangements for establishing diploma standards will vary from school district to school district, other institutions, agencies, or the like

will evaluate the adequacy of the evaluation and standard-setting procedures that led to particular students earning diplomas. Such "outsiders" are likely to raise questions about the evaluation process and level of standards employed. By doing a good job of documenting and explaining the procedures that were followed in designing evaluation methods and in establishing and applying appropriate standards, an institution can increase the acceptability or credibility of the diplomas that are granted.

The problem of demonstrating that high enough standards were indeed followed is a more difficult one. Critics are fond of pointing out examples of high school graduates who cannot read, college graduates who cannot write, or professional school graduates who cannot function effectively in the profession for which they were supposedly trained. The best an individual institution can probably do is to define clearly the planning, implementation, and monitoring procedures that it uses. The greater the degree of adherence to its policies that can be attained, and thus the greater the consistency in evaluation and standards application that is achieved, the less the justification for criticism.

Student Satisfaction

In addition to meeting the criteria cited above, an adequate evaluation and standard-setting arrangement should satisfy the majority of the students who have been evaluated. Provision needs to be made for periodic follow-up surveys of students who have gone through the competency assessment process. To what extent do the students feel that they were prepared adequately for the assessment procedures that were applied? Did the standards seem to be fair and realistic? Did they find that preparing for and experiencing the assessment was a valuable learning experience?

Was there an air of seriousness about the enterprise that was consistent with its degree of importance to the student's future educational program? Although students who do not meet standards initially can be expected to be disappointed, they should be able to recognize the reasonableness of decisions made. If they do not, further investigation of the adequacy of the system and its application would clearly be called for.

METHODS OF SETTING COMPETENCY STANDARDS

The growth of interest in setting competency standards for the high school diploma has caused many measurement professionals to examine past work in the area of standards setting. An article by John Meskauskas in the Winter (1976) issue of Review of Educational Research will be very helpful in this connection. Meskauskas (1976) reviewed a number of standards-setting models that have been applied in criterion-referenced testing situations where content areas have been tightly specified. He pointed out that different conceptions of the nature of "mastery" lead to different evaluation procedures, but that extensive practical applications are lacking for almost all methods. Another useful resource is a paper by Loretta Shepard, presented at last years NCME meeting (Shepard, 1976), which makes some practical recommendations on standard-setting approaches.

The discussion of methods of setting competency standards that follows is based on the assumption that three factors will contribute decisively to any school district's actions:

- The external requirements of other institutions, agencies and the like, e.g., state departments of education
- The basic educational mission and objectives of the school district
- The personal development goals of individual students

The weight given to any one of these factors will vary from situation to situation, but all will influence decisions about policy and practice in the area of competency standards for the diploma. There will always be interactions among the factors, and there will be times when serious conflicts will need to be resolved, as when a student's personal development goals do not mesh with state requirements or with any of the alternatives offered by a school district.

One basic determination that each institution will need to make concerns the number of domains within which it is prepared to set diploma standards and assessments. The larger the number of separate areas for which standards are set, the greater the proportion of students who will have difficulty meeting these standards. Clearly, critical decisions will need to be made regarding the respective roles of successful course completion and independent competency assessment in satisfying diploma requirements.

Once the areas to be assessed have been identified and possible assessment procedures selected, an approach to setting competency standards must be chosen. This section describes and evaluates seven possible methods:

- Development of Competence Scales
- Traditonal or Arbitrary Method
- Review of Test Questions or Tasks
- Identification of Prerequisite Skills
- Use of Normative Data
- Relationship to Course Performance
- Use of Preselected Groups

Development of Competence Scales

There has been a good deal of work on the setting of performance standards relating to training programs or licensing/certification settings

in which evaluators are checking for the presence or absence of well-defined skills. The evaluation of the broader educational objectives of high schools will frequently require more than a summing-of-units approach. In some instances it may prove more productive to develop competence scales that include:

- Definitions of a series of functional competence levels (usually 3 to 5)
- Descriptions of the kind of performance that should be associated with each level
- Illustrations (examples, models) of each level

The degree of elaboration of these functional competence levels would necessarily vary both as a function of the amount of student and faculty time that could be devoted to their development and as a function of the amount of available information on the qualities of student learning that should be dealt with in the scale. The initial step in developing a competence scale might be to list all the areas of knowledge and skill that should be considered. A next step could be to determine what level of knowledge or proficiency a person well above minimal competency might have. Such a step would provide the start for the high end of this competency scale. The other end might be developed by considering the absolute minimum skills that would be considered adequate. Through a combination of armchair speculation, expert review, and trial use, a competence scale could be developed that evaluators would use as a guide to judgments about students.

It would be a heroic enterprise to create a full spectrum of competence scales for every possible component of the high school experience. Institutions could move towards such a goal, though, by developing competence scales for very high priority objectives. It might be possible to draw on existing records of student performance and on the comments and experience of the staff and any outside evaluators. Use of a competence-scale approach over

over a period of years could result in the development of a file of models of scales with notes on their use that would help evaluators facing future scale-definition problems.

A partial sample of the type of competence scale that has been described can be found in the language proficiency scales developed by the U. S. Foreign Service Institute and used for many years with Peace Corps volunteers:

Elementary Proficiency

Speaking-1. Able to satisfy routine travel needs and minimum courtesy requirements. Can ask and answer questions on topics very familiar to him; within the scope of his very limited language experience can understand simple questions and statements, allowing for slowed speech, repetition or paraphrase; speaking vocabulary inadequate to express anything but the most elementary needs; errors in pronunciation and grammar are frequent, but can be understood by a native speaker used to dealing with foreigners attempting to speak his language; while topics which are "very familiar" and elementary needs vary considerably from individual to individual, any person at the S-1 level should be able to order a simple meal, ask for shelter or lodging, ask and give simple directions, make purchases, and tell time.

Reading-1. Able to read some personal and place names, street signs, office and shop designations, numbers, and isolated words and phrases. Can recognize all the letters in the printed version of an alphabetic system and high-frequency elements of a syllabary or a character system.

Limited Working Proficiency

Speaking-2. Able to satisfy routine social demands and limited work requirements. Can handle with confidence but not with facility most social situations including introductions and casual conversations about current events, as well as work, family, and autobiographical information; can handle limited work requirements, needing help in handling any complications or difficulties; can get the gist of most conversations on non-technical subjects (i.e., topics which require no specialized knowledge) and has a speaking vocabulary sufficient to express himself simply with some circumlocutions; accent, though often quite faulty, is intelligible; can usually handle elementary constructions quite accurately but does not have thorough or confident control of the grammar.

Reading-2. Able to read simple prose, in a form equivalent to typescript or printing, on subjects within a familiar context. With extensive use of a dictionary can get the general sense of routine business letters, international news items, or articles in technical fields within his competence.

Minimum Professional Proficiency

Speaking-3. Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social, and professional topics. Can discuss

particular interests and special fields of competence with reasonable ease; comprehension is quite complete for a normal rate of speech; vocabulary is broad enough that he rarely has to grope for a word; accent may be obviously foreign; control of grammar good; errors never interfere with understanding and rarely disturb the native speaker.

Reading-3. Able to read standard newspaper items addressed to the general reader, routine correspondence, reports and technical material in his special field. Can grasp the essentials of articles of the above types without using a dictionary; for accurate understanding moderately frequent use of a dictionary is required. Has occasional difficulty with unusually complex structures and low-frequency idioms.

Full Professional Proficiency

Speaking-4. Able to use the language fluently and accurately on all levels normally pertinent to professional needs. Can understand and participate in any conversation within the range of his experience with a high degree of fluency and precision of vocabulary; would rarely be taken for a native speaker, but can respond appropriately even in unfamiliar situations; errors of pronunciation and grammar quite rare; can handle informal interpreting from and into the language.

Reading-4. Able to read all styles and forms of the language pertinent to professional needs. With occasional use of a dictionary can read moderately difficult prose readily in any area directed to the general reader, and all material in his special field including official and professional documents and correspondence; can read reasonably legible handwriting without difficulty.

Native or Bilingual Proficiency

Speaking-5. Speaking proficiency equivalent to that of an educated native speaker. Has complete fluency in the language such that his speech on all levels is fully accepted by educated native speakers in all of its features, including breadth of vocabulary and idiom, colloquialisms, and pertinent cultural references.

Reading-5. Reading proficiency equivalent to that of an educated native. Can read extremely difficult and abstract prose, as well as highly colloquial writings and the classic literary forms of the language. With varying degrees of difficulty can read all normal kinds of handwritten documents.

Traditional or Arbitrary Method

When dealing with an examination or set of performance tasks, the worst of all possible methods of setting standards is to choose an arbitrary percentage of correct responses because that percentage has traditionally been associated with a "passing" score. The method has its roots in the percent grades which were used almost universally in schools and colleges until about 1920. Unfortunately, this approach (the "judgment uncontaminated by data" method) is the most frequently used process for setting standards.

The answer to the question "What seems right for a passing score?" usually represents a compromise between those who desire as high a percentage as possible to show their interest in quality education and those who are afraid that their standards will be impossible for the bulk of students to meet.

Clearly, no method that fails to take into account the nature of the learning experience, test questions or tasks, and the students' backgrounds is satisfactory. A standard such as 65 percent or 80 percent correct might be reasonable for certain tests for certain groups of students; it may be wildly inappropriate for other tests or sets of performance tasks with other students. As a rule of thumb, blanket performance standards should be avoided.

Review of Test Questions or Tasks

This method sets a standard for each set of questions or performance tasks used for a specific competency area. Informed judges examine each question or behavioral statement of a critical task and decide if a minimally competent student could answer the question correctly or perform the task adequately. The standard is simply the average number of questions or tasks that each judge has determined should be answered correctly or performed at a specified level of competence by minimally competent students.

One way to improve this procedure in a situation involving a test is to ask each judge to state the probability that a minimally competent person would answer each question correctly. The judges review the records of a number of minimally competent persons and estimate the proportions of such students who gave the correct answer. The sum of these proportions represents the minimally acceptable score. Where task review is involved, judges will

will need to decide on the weight to be given to different aspects such as the kind and degree of learning reflected in the produce or performance.

Identification of Prerequisite Skills

This method is clearly of most use when the student's attainment of an ultimate goal is unlikely unless certain prior "enabling" skills have been acquired. Success at advanced calculus, for example, is not likely unless the student has mastered elementary algebra. Similarly, success at discovering original proofs of mathematical theorems in abstract course work is very unlikely unless the student has previously (1) learned how to reproduce (in his own words) standard textbook proofs and (2) become thoroughly acquainted with several relatively concrete examples of the theory so as to have a strong intuitive feel for the abstractions. Although most courses of study do not enjoy the hierarchical precision of mathematics, the identification of the prerequisite skills method is most efficient at points where decisions are made concerning a student's ability to undertake a new unit in a hierarchical sequence of units.

This method uses pooled judgments, but the decisions are made on the basis of whether or not success on a question or task requires knowledge or skills that are essential to the future application of skills learned in the course or program. The method, in practice, is quite similar to that used in reviewing questions or tasks; it retains the same advantages and problems. But the focus is not on asking if a minimally competent person would get a particular question correct but rather if success on a question or task is required before some specified future activity is possible. A task analysis of the future activity can help to determine what behaviors are required in the performance of complicated tasks.

An example of the kind of analysis that can be particularly useful in setting competency requirements for the diploma was a survey of a national probability sample (N=5,096) of adults aged 16 and over. These adults were asked to describe the kinds of reading performed in the course of a normal day, the amount of time spent in carrying out the reading described, and the importance attached to the various kinds of reading performed. Results of this survey were published in Reading Activities of American Adults (Sharon, 1972).

Use of Normative Data

A norm-referenced test or a summary of the products and performances of a group of students can be used to make decisions about minimal competency levels. Even though this method has empirical elements, pooled judgment is still needed at some stage in the selection of a norming or comparison group. A group must be chosen which is sufficiently similar to the group of interest. In contrast to previously discussed methods in which judgments concerned questions or tasks, judgments in this method concern whether or not a student is just minimally competent. The assessment procedure is then applied to a group of such students. The mean score or average performance of that group may be taken as the minimum performance standard. Using normative data tends to insure the setting of a performance standard that is realistic in terms of the percentage of the population that is judged to have reached an adequate level of performance on the basis of the test. This method has the most ready application to situations in which an objectively scored examination or performance on a standard set of tasks will be the basis for decisions regarding a student's level of competence.

Relationship to Course Performance

This method can be applied to decisions about minimal competence when the skills required can be related to the content of a particular course. A test or other measurement procedure that has been selected or developed to be the basis for competence decisions is administered to students completing the course. A table or graph is prepared showing the relationship between test scores or performance ratings and grades earned in the regular course. A table can take the form of a grid containing a number of cells. Inside each cell is a number based on the number of people who had scores on the test and grades in the course corresponding to the position of the cell. Usually the data are organized to allow direct readings of probabilities of earning a particular course grade given a test score (e.g., of those who scored in the range 45-50 on the test, what percentage earned a grade of B or better?) Alternatively, scores on the examination could be related to course grades by means of a regression equation, which will give a predicted grade for any given test score. (Warren Willingham reviews possible methods of using course performance as a basis for setting a performance standard in a chapter entitled "Exemption from Requirements Already Mastered" in the book College Placement and Exemption.)

When the relationship-to-course-performance method is used, the choice of the test performance standard is made on the basis of the probability of obtaining various grades. Judgment is, of course, necessary in selecting an appropriate probability. One might, for example, set a performance standard so that there is a probability of .80 that the grade earned in a regular course would have been C (or some other level) or better.

Use of Preselected Groups

This method requires that information independent of the particular assessment procedure being applied be used to select one group of students who are above a school's minimum requirement and another that are not. Instructors' judgments should be used to divide students into the two groups. In making these judgments, it is essential to differentiate competencies from other factors that contribute to grades; e.g., punctuality, consistent attendance, classroom behavior, etc.

The assessment procedure is then applied to both groups, and the distribution of scores obtained. A competency standard is then selected that best discriminates between the two groups.

This method has the important advantage of inherent concurrent validation: The test is validated against independent informed judgments of student competency. Furthermore, this procedure allows the use of classical measures of test quality such as reliability, validity, item discrimination, and the like.

Error of Measurement

The methods of setting competency standards that are described in this paper combine human judgment and empirical data in a systematic process. Regardless of the quality of the procedure, however, some error of measurement will result. This "error" is not attributable to arithmetical mistakes that could be avoided by more careful calculations. Such error is related to the complexity and variability of people and the difficulty of measuring their skills and knowledge. If you were somehow able to achieve complete knowledge of each student's competencies and had a perfect standard-setting procedure, you would be able to detect some degree of error in any operational system.

There are two major types of errors possible in setting standards. If

the standards are set too high, the probability increases that those who should be considered competent will be classified as not competent. If the standards are set too low, the probability increases that those who should fail will be passed. Clearly, the setting of standards must take into account the costs or losses associated with misclassification of each type. In every situation, basic principles of fairness, realism, and consistency will apply, but it will also be necessary to consider the social implications of decisions about students' competencies. An analogy might be made to testing applicants for drivers' licenses. There an incorrect judgment that an applicant has acquired certain important competencies might have the effect of enabling the individual to avoid some needed additional training, allowing an incompetent driver to emerge as a hazard to others.

A Final Thought

The job of setting appropriate performance standards is as critical as it is difficult. Developers of evaluation procedures as part of diploma requirements should allow sufficient time in their projects for a thoughtful analysis of the possible approaches. If at all possible, more than one method of setting standards should then be tried, and the results compared and evaluated. After this evaluation, a decision can be made regarding the method of methods that will be employed for each assessment procedure.

Acknowledgment

I would like to acknowledge the contribution of my colleague Michael Zieky of Educational Testing Service to the last section of this paper. My treatment of standards setting methods draws heavily on analyses he carried out for workshops on criterion-referenced and objectives-referenced measurement.

REFERENCES

Baum, Joan The Politics of Back to Basics. Change, Vol. 8, 10, November 1976, pp. 32-36.

Clark, James P. and Thomson, Scott D. Competency tests and graduation requirements. Reston, Virginia: National Association of Secondary School Principals, 1976.

Keeton, Morris Reconnoitering for a CAEL strategy to improve the assessment of the outcome of experiential learning, in Morris Keeton (Ed.), Reflections on experiential learning and its uses. CAEL Working Paper No. 3. Princeton, N.J.: Educational Testing Service, 1974, pp. 1-11.

Meskauskas, John A. Evaluation models for criterion-referenced testing: Views regarding mastery and standard-setting. Review of Educational Research, 1976, 46, 133-158.

National Education Association Changing high school requirements, Briefing Memo, 11, June 1976.

Sharon, Amiel T. Reading activities of American adults. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois, April 1972.

Shepard, Loretta A. Setting standards and living with them. Paper presented at the annual meeting of National Council on Measurement in Education, San Francisco, April 1976.

Wilds, Claudia P. The oral interview test, in Randall L. Jones and Bernard Spolsky (Eds.), Testing language proficiency. Arlington, VA.: Center for Applied Linguistics, 1975.

Willingham, Warren W. College placement and exemption. New York: College Entrance Examination Board, 1974.